Supplementary material, Appendix 1: Proofs

Proof of Theorem 1

The first partial derivatives of $\log[s(y|\mu,\sigma,\gamma)]$ are given by

$$\begin{aligned}
\frac{\partial}{\partial\mu}\log[s(y|\mu,\sigma_1,\sigma_2)] &= -\frac{1}{\sigma_1}\frac{f'\left(\frac{y-\mu}{\sigma_1}\right)}{f\left(\frac{y-\mu}{\sigma_1}\right)}I_{(-\infty,\mu)}(y)-\frac{1}{\sigma_2}\frac{f'\left(\frac{y-\mu}{\sigma_2}\right)}{f\left(\frac{y-\mu}{\sigma_2}\right)}I_{[\mu,\infty)}(y),\\
\frac{\partial}{\partial\sigma_1}\log[s(y|\mu,\sigma_1,\sigma_2)] &= -\frac{1}{\sigma_1+\sigma_2}-\frac{y-\mu}{\sigma_1^2}\frac{f'\left(\frac{y-\mu}{\sigma_1}\right)}{f\left(\frac{y-\mu}{\sigma_1}\right)}I_{(-\infty,\mu)}(y),\\
\frac{\partial}{\partial\sigma_2}\log[s(y|\mu,\sigma_1,\sigma_2)] &= -\frac{1}{\sigma_1+\sigma_2}-\frac{y-\mu}{\sigma_2^2}\frac{f'\left(\frac{y-\mu}{\sigma_2}\right)}{f\left(\frac{y-\mu}{\sigma_2}\right)}I_{[\mu,\infty)}(y).
\end{aligned}$$

Then the entries of the Fisher information matrix of (μ,σ_1,σ_2) are given by

$$\begin{aligned}
I_{11} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\mu}\log[s(y|\mu,\sigma_1,\sigma_2)]\right)^2\right]=\frac{2\alpha_1}{\sigma_1\sigma_2},\\
I_{22} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\sigma_1}\log[s(y|\mu,\sigma_1,\sigma_2)]\right)^2\right]=\frac{\alpha_2}{\sigma_1(\sigma_1+\sigma_2)}+\frac{\sigma_2}{\sigma_1(\sigma_1+\sigma_2)^2},\\
I_{33} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\sigma_2}\log[s(y|\mu,\sigma_1,\sigma_2)]\right)^2\right]=\frac{\alpha_2}{\sigma_2(\sigma_1+\sigma_2)}+\frac{\sigma_1}{\sigma_2(\sigma_1+\sigma_2)^2},\\
I_{12} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\mu}\log[s(y|\mu,\sigma_1,\sigma_2)]\right)\left(\frac{\partial}{\partial\sigma_1}\log[s(y|\mu,\sigma_1,\sigma_2)\right)\right]=-\frac{2\alpha_3}{\sigma_1(\sigma_1+\sigma_2)},\\
I_{13} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\mu}\log[s(y|\mu,\sigma_1,\sigma_2)]\right)\left(\frac{\partial}{\partial\sigma_2}\log[s(y|\mu,\sigma_1,\sigma_2)\right)\right]=\frac{2\alpha_3}{\sigma_2(\sigma_1+\sigma_2)},\\
I_{23} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\sigma_1}\log[s(y|\mu,\sigma_1,\sigma_2)]\right)\left(\frac{\partial}{\partial\sigma_2}\log[s(y|\mu,\sigma_1,\sigma_2)\right)\right]=-\frac{1}{(\sigma_1+\sigma_2)^2}.
\end{aligned}$$

$\square$

Proof of Theorem 2

The determinant of the Fisher information matrix is

$$|I(\mu,\sigma_1,\sigma_2)|=\frac{2\alpha_2\left(\alpha_1+\alpha_1\alpha_2-2\alpha_3^2\right)}{\sigma_1^2\sigma_2^2(\sigma_1+\sigma_2)^2}.$$

We will first prove that $\alpha_2>0$. From the definition of α_2 it can only be zero if $1+tf'(t)/f(t)=0$ whenever $f(t)>0$. This means that $f(t)=-tf'(t)$ and this only happens if

$f(t) = K/t$ for any positive K. The latter, however, is not a probability density function on $\mathbb{R}$. Thus, α_2 can not be zero.

Next, we will prove that $\alpha_1(1+\alpha_2) > 2\alpha_3^2$. Applying the Cauchy-Schwarz inequality we have $\alpha_1(1+\alpha_2) \geq 2\alpha_3^2$. We will show that this is a strict inequality. The condition in Theorem 2 implies that

$$0 < \int_0^\infty t \left[\frac{f'(t)}{f(t)}\right]^2 f(t)\,dt.$$

Let

$$\phi(t) = \left|\frac{f'(t)}{\sqrt{f(t)}}\right| > 0 \text{ a.e. and } \psi(t) = t\left|\frac{f'(t)}{\sqrt{f(t)}}\right| > 0 \text{ a.e.}$$

Note that $[\beta\phi(t) + \psi(t)]^2 > 0$ a.e. for any $\beta \in \mathbb{R}$, and thus

$$0 < \int_0^\infty [\beta\phi(t) + \psi(t)]^2\,dt = \beta^2 \int_0^\infty \phi^2(t)\,dt + 2\beta \int_0^\infty \phi(t)\psi(t)\,dt + \int_0^\infty \psi^2(t)\,dt.$$

This is a polynomial of degree 2 in β with positive coefficients and no real roots, implying that the discriminant is negative, so that

$$\left[\int_0^\infty t\left[\frac{f'(t)}{f(t)}\right]^2 f(t)\,dt\right]^2 < \left[\int_0^\infty t^2\left[\frac{f'(t)}{f(t)}\right]^2 f(t)\,dt\right]\left[\int_0^\infty \left[\frac{f'(t)}{f(t)}\right]^2 f(t)\,dt\right].$$

□

Proof of Theorem 3

The first partial derivatives of $\log[s(y|\mu,\sigma,\gamma)]$ are given by

$$\begin{aligned}
\frac{\partial}{\partial\mu}\log[s(y|\mu,\sigma,\gamma)] &= -\frac{1}{\sigma b(\gamma)}\frac{f'\left(\frac{y-\mu}{\sigma b(\gamma)}\right)}{f\left(\frac{y-\mu}{\sigma b(\gamma)}\right)} I_{(-\infty,\mu)}(y) - \frac{1}{\sigma a(\gamma)}\frac{f'\left(\frac{y-\mu}{\sigma a(\gamma)}\right)}{f\left(\frac{y-\mu}{\sigma a(\gamma)}\right)} I_{[\mu,\infty)}(y),\\
\frac{\partial}{\partial\sigma}\log[s(y|\mu,\sigma,\gamma)] &= -\frac{1}{\sigma} - \frac{y-\mu}{\sigma^2 b(\gamma)}\frac{f'\left(\frac{y-\mu}{\sigma b(\gamma)}\right)}{f\left(\frac{y-\mu}{\sigma b(\gamma)}\right)} I_{(-\infty,\mu)}(y) - \frac{y-\mu}{\sigma^2 a(\gamma)}\frac{f'\left(\frac{y-\mu}{\sigma a(\gamma)}\right)}{f\left(\frac{y-\mu}{\sigma a(\gamma)}\right)} I_{[\mu,\infty)}(y),\\
\frac{\partial}{\partial\gamma}\log[s(y|\mu,\sigma,\gamma)] &= -\frac{a'(\gamma)+b'(\gamma)}{a(\gamma)+b(\gamma)} - \frac{y-\mu}{\sigma}\frac{b'(\gamma)}{b(\gamma)^2}\frac{f'\left(\frac{y-\mu}{\sigma b(\gamma)}\right)}{f\left(\frac{y-\mu}{\sigma b(\gamma)}\right)} I_{(-\infty,\mu)}(y)\\
&- \frac{y-\mu}{\sigma}\frac{a'(\gamma)}{a(\gamma)^2}\frac{f'\left(\frac{y-\mu}{\sigma a(\gamma)}\right)}{f\left(\frac{y-\mu}{\sigma a(\gamma)}\right)} I_{[\mu,\infty)}(y).
\end{aligned}$$

Thus, the entries of the Fisher information matrix of (μ, σ, γ) are

$$\begin{aligned}
I_{11} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\mu}\log[s(y|\mu,\sigma,\gamma)]\right)^2\right] = \frac{2\alpha_1}{a(\gamma)b(\gamma)\sigma^2},\\
I_{22} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\sigma}\log[s(y|\mu,\sigma,\gamma)]\right)^2\right] = \frac{\alpha_2}{\sigma^2},\\
I_{33} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\gamma}\log[s(y|\mu,\sigma,\gamma)]\right)^2\right] = \frac{\alpha_2+1}{a(\gamma)+b(\gamma)}\left[\frac{b'(\gamma)^2}{b(\gamma)}+\frac{a'(\gamma)^2}{a(\gamma)}\right] - \left[\frac{a'(\gamma)+b'(\gamma)}{a(\gamma)+b(\gamma)}\right]^2,\\
I_{12} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\mu}\log[s(y|\mu,\sigma,\gamma)]\right)\left(\frac{\partial}{\partial\sigma}\log[s(y|\mu,\sigma,\gamma)\right)\right] = 0,\\
I_{13} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\mu}\log[s(y|\mu,\sigma,\gamma)]\right)\left(\frac{\partial}{\partial\gamma}\log[s(y|\mu,\sigma,\gamma)\right)\right]\\
&= \frac{2\alpha_3}{\sigma[a(\gamma)+b(\gamma)]}\left[\frac{a'(\gamma)}{a(\gamma)}-\frac{b'(\gamma)}{b(\gamma)}\right],\\
I_{23} &= \mathbb{E}\left[\left(\frac{\partial}{\partial\sigma}\log[s(y|\mu,\sigma,\gamma)]\right)\left(\frac{\partial}{\partial\gamma}\log[s(y|\mu,\sigma,\gamma)\right)\right]\\
&= \frac{\alpha_2}{\sigma}\left[\frac{a'(\gamma)+b'(\gamma)}{a(\gamma)+b(\gamma)}\right].
\end{aligned}$$

□

Proof of Theorem 4

Note that

$$\frac{d}{d\gamma}AG(\gamma) = 2\frac{a'(\gamma)b(\gamma)-a(\gamma)b'(\gamma)}{[a(\gamma)+b(\gamma)]^2} = 2\frac{a(\gamma)b(\gamma)\lambda(\gamma)}{[a(\gamma)+b(\gamma)]^2},$$

so that

$$\frac{dAG(\gamma)}{d\gamma} > 0 \Leftrightarrow \lambda(\gamma) > 0 \text{ and } \frac{dAG(\gamma)}{d\gamma} < 0 \Leftrightarrow \lambda(\gamma) < 0.$$

□

Proof of Theorem 5

First of all, consider the independence Jeffreys prior (6) and the change of variable (7), then

$$\begin{aligned}
\pi_I(\mu,\sigma,\gamma) &\propto \frac{|a'(\gamma)b(\gamma)-a(\gamma)b'(\gamma)|\sqrt{[b(\gamma)+\alpha_2[a(\gamma)+b(\gamma)]][a(\gamma)+\alpha_2[a(\gamma)+b(\gamma)]]}}{\sigma\sqrt{a(\gamma)b(\gamma)}[a(\gamma)+b(\gamma)]^2}\\
&\leq \frac{(\alpha_2+1)|a'(\gamma)b(\gamma)-a(\gamma)b'(\gamma)|}{\sigma\sqrt{a(\gamma)b(\gamma)}[a(\gamma)+b(\gamma)]}.
\end{aligned}$$

For the particular choice $\{a(\gamma), b(\gamma)\} = \{\gamma, 1/\gamma\}$, the upper bound of $\pi_I(\mu, \sigma, \gamma)$ is proportional to $[\sigma(1+\gamma^2)]^{-1}$. Now, the proof of (i) and (ii) is as follows.

(i) Applying Theorem 1 from Fernández and Steel(1999) and using this upper bound we can derive the properness of the posterior distribution of (μ, σ, γ). Now, since the mapping $(\mu, \sigma, \gamma) \leftrightarrow (\mu, \sigma_1, \sigma_2)$ is one-to-one, it follows that the posterior distribution of $(\mu, \sigma_1, \sigma_2)$ is proper.

(ii) The proof follows analogously by applying Theorem 2 from Fernández and Steel(1999). □

Proof of Theorem 6

Let f be a scale mixture of normals with τ_j the mixing variable associated with y_j and where the τ_js are independent random variables defined on $\mathbb{R}^+$ with distribution P_{τ_j}.

(i) Integrating with respect to μ over a subspace we get a lower bound for the marginal distribution of $(y_1, ..., y_n)$ which is proportional to

$$\int_{\mathbb{R}_+^n} \int_\Gamma \int_0^\infty \int_{-\infty}^{y_{(1)}} \left(\prod_{j=1}^n \tau_j^{\frac{1}{2}} \right) \frac{\sigma^{-(n+1)}}{[a(\gamma)+b(\gamma)]^n} \exp\left[-\frac{1}{2\sigma^2 a(\gamma)^2} \sum_{j=1}^n \tau_j (y_j - \mu)^2 \right]$$
$$\times \quad \pi(\gamma)\, d\mu d\sigma d\gamma dP_{(\tau_1,\ldots,\tau_n)}.$$

Using the change of variable $\vartheta = \sigma a(\gamma)$, we can rewrite the lower bound as follows

$$\int_\Gamma \left[\frac{a(\gamma)}{a(\gamma)+b(\gamma)} \right]^n \pi(\gamma)\, d\gamma \int_{\mathbb{R}_+^n} \int_0^\infty \int_{-\infty}^{y_{(1)}} \left(\prod_{j=1}^n \tau_j^{\frac{1}{2}} \right) \vartheta^{-(n+1)}$$
$$\times \quad \exp\left[-\frac{1}{2\vartheta^2} \sum_{j=1}^n \tau_j (y_j - \mu)^2 \right] d\mu d\vartheta dP_{(\tau_1,\ldots,\tau_n)},$$

and the result follows.

(ii) We can get an upper bound for the marginal distribution of $(y_1, ..., y_n)$ proportional to

$$\int_{\mathbb{R}_n^+} \int_\Gamma \int_0^\infty \int_{-\infty}^{\infty} \left(\prod_{j=1}^n \tau_j^{\frac{1}{2}} \right) \frac{\sigma^{-(n+1)}}{[a(\gamma)+b(\gamma)]^n} \exp\left[-\frac{1}{2\sigma^2 h(\gamma)^2} \sum_{j=1}^n \tau_j (y_j - \mu)^2 \right]$$
$$\times \quad \pi(\gamma)\, d\mu d\sigma d\gamma dP_{(\tau_1,\ldots,\tau_n)},$$

where $h(\gamma) = \max\{a(\gamma), b(\gamma)\}$. Consider the change of variable $\vartheta = \sigma h(\gamma)$ and rewrite the upper bound as follows

$$\int_\Gamma \left[\frac{h(\gamma)}{a(\gamma)+b(\gamma)}\right]^n \pi(\gamma)\, d\gamma \int_{\mathbb{R}_n^+} \int_0^\infty \int_{-\infty}^\infty \left(\prod_{j=1}^n \tau_j^{\frac{1}{2}}\right) \vartheta^{-(n+1)}$$
$$\times \quad \exp\left[-\frac{1}{2\vartheta^2}\sum_{j=1}^n \tau_j (y_j-\mu)^2\right] d\mu d\vartheta dP_{(\tau_1,\dots,\tau_n)}.$$

Fernández and Steel (2000, Th. 1) show that the integral in $\mu, \vartheta, \tau_1, ..., \tau_n$ is finite if $n \geq 2$. Then, by Theorem 1 from Fernández and Steel(1999), the existence of the integral in γ is a sufficient condition for the properness of the posterior distribution of (μ, σ, γ). The result then follows from

$$\int_\Gamma \left[\frac{h(\gamma)}{a(\gamma)+b(\gamma)}\right]^n \pi(\gamma)\, d\gamma \leq \int_\Gamma \pi(\gamma)\, d\gamma.$$

(iii) The proof follows analogously by applying Theorem 2 from Fernández and Steel(1999).

□

Proof of Theorem 7

If f is a scale mixture of normals, then integrating over a subspace with respect to μ we get a lower bound for the marginal distribution of $(y_1, ..., y_n)$ which is proportional to

$$\int_{\mathbb{R}_+^n} \int_\Gamma \int_0^\infty \int_{-\infty}^{y_{(1)}} \left(\prod_{j=1}^n \tau_j^{\frac{1}{2}}\right) \frac{\sigma^{-(n+2)}}{[a(\gamma)+b(\gamma)]^n} \exp\left[-\frac{1}{2\sigma^2 a(\gamma)^2}\sum_{j=1}^n \tau_j (y_j-\mu)^2\right]$$
$$\times \frac{|\lambda(\gamma)|}{a(\gamma)+b(\gamma)}\, d\mu d\sigma d\gamma dP_{(\tau_1,\dots,\tau_n)}.$$

Using the change of variable $\vartheta = \sigma a(\gamma)$, we can rewrite this lower bound as follows

$$\int_\Gamma \left[\frac{a(\gamma)}{a(\gamma)+b(\gamma)}\right]^{n+1} |\lambda(\gamma)|\, d\gamma \int_{\mathbb{R}_+^n} \int_0^\infty \int_{-\infty}^{y_{(1)}} \left(\prod_{j=1}^n \tau_j^{\frac{1}{2}}\right) \vartheta^{-(n+2)}$$
$$\times \exp\left[-\frac{1}{2\vartheta^2}\sum_{j=1}^n \tau_j (y_j-\mu)^2\right] d\mu d\vartheta dP_{(\tau_1,\dots,\tau_n)}.$$

Therefore, the existence of the first integral is a necessary condition for the properness of the posterior distribution of (μ, σ, γ). □

Proof of Theorem 8

The proof of (i) is as follows. If f is normal, defining $h(\gamma) = \max\{a(\gamma), b(\gamma)\}$ we get an upper bound for the marginal distribution of $(y_1, ..., y_n)$ which is proportional to

$$\int_{-\infty}^{\infty}\int_{\Gamma}\int_{0}^{\infty} \frac{\pi_J(\mu,\sigma,\gamma)}{[a(\gamma)+b(\gamma)]^n\sigma^n} \exp\left[-\frac{1}{2\sigma^2 h(\gamma)^2}\sum_{j=1}^{n}(y_j-\mu)^2\right] d\sigma d\gamma d\mu$$
$$\propto \int_{-\infty}^{\infty}\left[\sum_{j=1}^{n}(y_j-\mu)^2\right]^{-\frac{n+1}{2}} d\mu \int_{\Gamma} \frac{h(\gamma)^{n+1}}{[a(\gamma)+b(\gamma)]^{n+1}}|\lambda(\gamma)|d\gamma.$$

The first integral exists if $n \geq 2$ and at least 2 observations are different. Then the existence of the second integral is a sufficient condition for the existence of the posterior distribution. For the second integral we use that

$$\int_{\Gamma} \frac{h(\gamma)^{n+1}}{[a(\gamma)+b(\gamma)]^{n+1}}|\lambda(\gamma)|d\gamma \leq \int_{\Gamma}|\lambda(\gamma)|d\gamma,$$

which is finite by assumption. If f is Laplace, analogously to the normal case we get an upper bound for the marginal distribution of $(y_1, ..., y_n)$ which is proportional to

$$\int_{-\infty}^{\infty}\int_{\Gamma}\int_{0}^{\infty} \frac{\pi_J(\mu,\sigma,\gamma)}{[a(\gamma)+b(\gamma)]^n\sigma^n} \exp\left[-\frac{1}{\sigma h(\gamma)}\sum_{j=1}^{n}|y_j-\mu|\right] d\sigma d\gamma d\mu$$
$$\propto \int_{-\infty}^{\infty}\left[\sum_{j=1}^{n}|y_j-\mu|\right]^{-(n+1)} d\mu \int_{\Gamma} \frac{h(\gamma)^{n+1}}{[a(\gamma)+b(\gamma)]^{n+1}}|\lambda(\gamma)|d\gamma,$$

and the same argument leads to the result.

Result (ii) follows immediately from Corollary 6.

For (iii) let us assume, without loss of generality, that $AG(\gamma)$ is an increasing function and $\Gamma = (\underline{\gamma}, \overline{\gamma})$. First, note that we can rewrite $AG(\gamma)$ as follows

$$AG(\gamma) = \tanh\left\{\frac{1}{2}\log\left[\frac{a(\gamma)}{b(\gamma)}\right]\right\}.$$

Then

$$\lim_{\gamma\to\overline{\gamma}} AG(\gamma) = 1 \quad\Leftrightarrow\quad \lim_{\gamma\to\overline{\gamma}} \log\left[\frac{a(\gamma)}{b(\gamma)}\right] = \infty$$
$$\lim_{\gamma\to\underline{\gamma}} AG(\gamma) = -1 \quad\Leftrightarrow\quad \lim_{\gamma\to\underline{\gamma}} \log\left[\frac{a(\gamma)}{b(\gamma)}\right] = -\infty,$$

which contradicts the assumption that $\lambda(\gamma)$ is absolutely integrable. The result is analogous if AG is decreasing. $\square$

Proof of Theorem 9

From Theorem 6(ii) and (iii) we know that properness of $\pi(\gamma)$ in (23) is sufficient for existence of the posterior. The AG beta prior implies a proper prior for AG when $\alpha_0, \beta_0 > 0$. From Theorem 4 the condition that $\lambda(\gamma)$ does not change sign is equivalent to AG being a one-to-one transformation of γ. Thus, the induced prior on γ will be proper and the result follows. $\square$

Supplementary material, Appendix 2: Simulation Study

In this section we investigate the empirical coverage of the 95% posterior credible intervals, defined by the 2.5th and 97.5th percentiles. We simulate $N = 10,000$ data sets of size $n = 30, 100$ and 1000 from seven sampling models, Models 1-5 described in Section 4 plus two additional models described below, where we take f to be a normal distribution throughout, and analyse these data using the corresponding Bayesian model. Model 7 corresponds to the Logistic AG model model with AG beta prior and $\alpha_0 = \beta_0 = 1$, and Model 8 consists of the Inverse scale factors model with AG beta prior and $\alpha_0 = \beta_0 = 1$. For each of these N datasets, a sample of size $3,000$ was obtained from the posterior distribution using a Markov chain Monte Carlo sampler after a burn-in period of $5,000$ iterations and thinned to every 50th iteration. Finally, the proportion of 95% credible intervals that include the true value of the parameter was calculated. Results are presented in Tables 1-7. For Model 3 we know that the truncation to a finite interval is what makes the posterior well-defined. To investigate how sensitive the results are to the particular value chosen for B, we have experimented with various values. Models 5, 7 and 8 employ the same sort of prior with different parameterizations of the sampling model (9), while Models 1–4 differ in both the kind of prior employed and the parameterization of the sampling model.

Sample size	$n = 30$		$n = 100$		$n = 1000$	
Parameters	$\sigma_1 = 2.0$ $\sigma_2 = 0.5$	$\sigma_1 = 0.66$ $\sigma_2 = 1.50$	$\sigma_1 = 2.0$ $\sigma_2 = 0.5$	$\sigma_1 = 0.66$ $\sigma_2 = 1.50$	$\sigma_1 = 2.0$ $\sigma_2 = 0.5$	$\sigma_1 = 0.66$ $\sigma_2 = 1.50$
μ	0.976	0.967	0.971	0.956	0.948	0.953
σ_1	0.961	0.951	0.974	0.958	0.947	0.949
σ_2	0.975	0.971	0.961	0.951	0.948	0.950

Table 1: Coverage proportions. Two-piece model in (2) with independence Jeffreys prior (Model 1)

Sample size	$n = 30$		$n = 100$		$n = 1000$	
Parameter	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$
μ	0.971	0.967	0.954	0.955	0.947	0.948
σ	0.959	0.960	0.947	0.945	0.953	0.954
γ	0.971	0.969	0.957	0.957	0.948	0.952

Table 2: Coverage proportions. ϵ-skew model with independence Jeffreys prior (Model 2)

Sample size	$n = 30$		$n = 100$		$n = 1000$	
Parameter	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$
μ	0.967	0.964	0.949	0.953	0.948	0.949
σ	0.995	0.991	0.952	0.960	0.948	0.947
γ	0.964	0.965	0.949	0.952	0.948	0.947

Table 3: Coverage proportions. Logistic AG model with Jeffreys prior (Model 3) and $B = 3$

Size	$n = 30$		$n = 100$		$n = 1000$	
Parameter	$\gamma = 0.5$	$\gamma = 1.5$	$\gamma = 0.5$	$\gamma = 1.5$	$\gamma = 0.5$	$\gamma = 1.5$
μ	0.969	0.967	0.963	0.950	0.949	0.946
σ	0.992	0.972	0.965	0.949	0.947	0.949
γ	0.967	0.971	0.967	0.950	0.950	0.948

Table 4: Coverage proportions: Inverse scale factors model with modified Jeffreys prior (Model 4)

Size	$n = 30$		$n = 100$		$n = 1000$	
Parameter	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$
μ	0.968	0.967	0.960	0.959	0.947	0.951
σ	0.994	0.993	0.968	0.970	0.947	0.951
γ	0.968	0.969	0.964	0.964	0.948	0.950

Table 5: Coverage proportions: ϵ-skew model with AG beta prior and $\alpha_0 = \beta_0 = 1$ (Model 5).

Size	$n = 30$		$n = 100$		$n = 1000$	
Parameter	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$
μ	0.965	0.965	0.956	0.961	0.945	0.950
σ	0.992	0.994	0.964	0.966	0.950	0.952
γ	0.968	0.968	0.960	0.965	0.947	0.948

Table 6: Coverage proportions: Logistic AG model with AG beta prior and $\alpha_0 = \beta_0 = 1$ (Model 7).

Size	$n = 30$		$n = 100$		$n = 1000$	
Parameter	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$	$\gamma = 0.5$	$\gamma = -0.5$
μ	0.969	0.973	0.952	0.949	0.949	0.948
σ	0.986	0.973	0.963	0.953	0.950	0.951
γ	0.968	0.976	0.959	0.951	0.946	0.951

Table 7: Coverage proportions: Inverse scale factors model with AG beta prior and $\alpha_0 = \beta_0 = 1$ (Model 8).

All models lead to coverage probabilities above the nominal level for samples of size $n = 30$, especially in the case of σ for Models 3–5 and 7. Once we increase the sample size to $n = 100$, the coverage is quite close to the nominal value, except for one setting for Model 1, where the coverage is still a bit high. As we further increase to samples of 1000 observations, all cases lead to coverage very close to 95%, as we would expect. The simulation standard errors are around 0.002 for all cases, so that for large n most differences in the tables can simply be accounted for by Monte Carlo error. For Model 3, the choice of B (we have also tried $B = 10$ and $B = 30$) did not seem to have any noticeable effect. Comparing Tables 2 and 5, Tables 3 and 6 and Tables 4 and 7 allows us to assess the difference in coverage between the AG beta prior and the other priors, and we can conclude these differences are quite small. The only exception is the performance for σ with 30 observations from the ϵ-skew model, where the independence Jeffreys prior leads to better coverage. Overall, the frequentist coverage properties of the models examined are pretty good, with perhaps Model 2 displaying the best performance.

We also conducted the same simulation study using a skewed version of a Student-t sampling model with 2 degrees of freedom and we observed a rather similar behaviour of the coverage proportions. Interestingly, however, the coverage for the ϵ-skew model with $n = 30$ is better in this case with the AG beta prior than under the independence Jeffreys and the overall coverage for σ in small samples is better than with the skewed normal throughout.